

Amendments to the Specification:

Please replace the first full paragraph on page 4 with the following amended paragraph:

These ~~feature~~ features each has advantages including a low-bit-rate hybrid coder using the voicing of weakly-voiced frames to enhance the waveform coder and avoiding phase discontinuities at the switching between parametric and waveform coded frames.

Please replace the last paragraph on page 6 with the following amended paragraph:

Analysis by Synthesis 114 is used by the CELP coder for weakly-voiced frames to encode the pitch, pitch-predictor gain, fixed-codebook contribution, and codebook gain. The initial pitch estimate is obtained from the pitch-and-voicing analysis. The fixed codebook is a sparse codebook with four pulses per 10 ms (80-sample) sub-frame. The pitch-predictor gain and the fixed excitation gain are quantized jointly by Quantization 442 110.

Please replace the fifth paragraph on page 7 with the following amended paragraph:

Postfilter 144 with coefficients derived from LP parameters provides ~~...~~ enhanced formant peaks.

Please replace the first paragraph on page 12 with the following amended paragraph:

The  $\{X[k]\}$  may be estimated by applying a discrete Fourier transform to the samples of a single period (or small number of periods) of  $e(n)$  as in Figures ~~3b-3e~~ 2a-2b. The preferred embodiment only uses the magnitudes of the Fourier coefficients, although the phases could also be used. Because the LP residual components  $\{e(n)\}$  are real, the discrete Fourier transform coefficients  $\{X(k)\}$  are conjugate symmetric:  $X(k) = X^*(N-k)$  for an  $N$ -point discrete Fourier transform. Thus only half of the  $\{X(k)\}$  need be used for magnitude considerations. Of course, with a pitch period of  $p$  samples,  $N$  will be an integer equal to  $[p]$  or  $[p]+1$ .

Please replace the first paragraph on page 12 with the following amended paragraph:

(2) apply speech activity detection to each of the ~~six~~ eight 20-sample sub-frames of the frame; the speech activity detection may be by the sum of squares of samples with a threshold.

Please replace paragraph (12) on page 15 with the following amended paragraph:

(12) pick the pitch candidate as follows (compare Figure 3): if  $pcorr[0]$  is less than  $4*threshold$ , then put  $i = -1$ ; if  $pcorr[0]$  is at least  $4*threshold$ , then  $i = 0$  unless  $pcorr[k]$  is at least  $0.8*pcorr[0]$ , then take  $i =$  the largest such  $k$  unless additionally  $pcorr[k]$  is less than  $0.9*pcorr[0]$  in which case take  $i = -1$ .